

# JBrowse

PAG 2019

Scott Cain

Ontario Institute for Cancer Research

GMOD Coordinator

WormBase Developer

[scott@scottcain.net](mailto:scott@scottcain.net)



# What is GMOD?

- A set of interoperable open-source **software** components for visualizing, annotating, and managing biological data.
- An active **community** of developers and users asking diverse questions, and facing common challenges, with their biological data.



# Who uses GMOD?





# Other GMOD Projects

- GBrowse - “old school” Genome Browser, based on BioPerl
- MAKER - Tool for doing whole genome *de novo* annotation
- Chado - Organism-agnostic database schema for genomics and related data
- Apollo - Web-based feature annotation editor based on JBrowse
- Tripal - Web framework based on Drupal/PHP; provides user interface to data in Chado (Right behind me at 4:00!)
- InterMine - Very powerful query engine/web UI for biological data (loaders for GFF, Chado)
- Galaxy - Very powerful workflow editor—lets you create and easily rerun complex workflows.

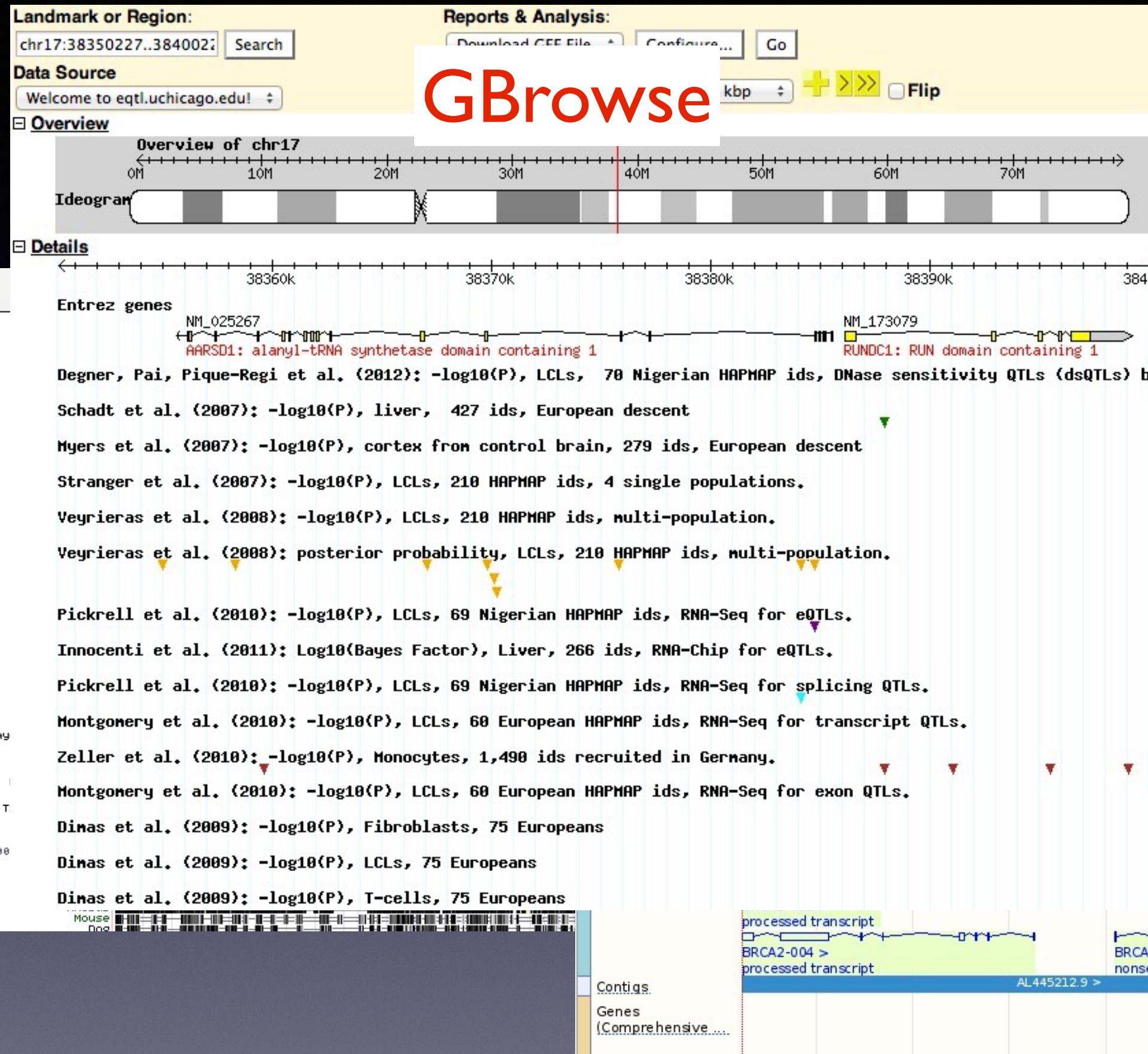


# What are Genome Browsers good for?

- Visualizing dense data from a multitude of sources (genes from a GFF file, RNASeq data from a BAM file, variation data from a VCF file)
- Homology and gene expression support for gene models



# GBrowse



ers

Ensembl

- Many “specialty” browsers (eg, Biodalliance, Savant)



# Why Install Your Own?

- You have data no one else has
- You want to be able to share it with your group, community, the world (potentially with “less savvy” users)
- You want to have control over how it looks



# Why JBrowse?

- (Fairly) Easy install
- Good user experience (getting close to a browser-desktop hybrid)
- Good community support (mailing lists, tutorials, software updates)



# Installation

- Only requires:
  - Web server (apache, lighttpd, nginx, etc)
  - Perl/make/standard unix-y tools
  - Really easy to get via AWS (Docker too, but we won't be using that)



# JBrowse Attributes

- Do everything possible on the client side, in JavaScript.
- Fast, smooth navigation.
- Supports GFF3, BED, Bio::DB::\*, Chado, WIG, BAM, BigWig, VCF, and UCSC import (intron/exon structure, name lookups, quantitative plots).
- Is stably funded by NHGRI.
- Is open source, of course.
- Did I mention it's fast?



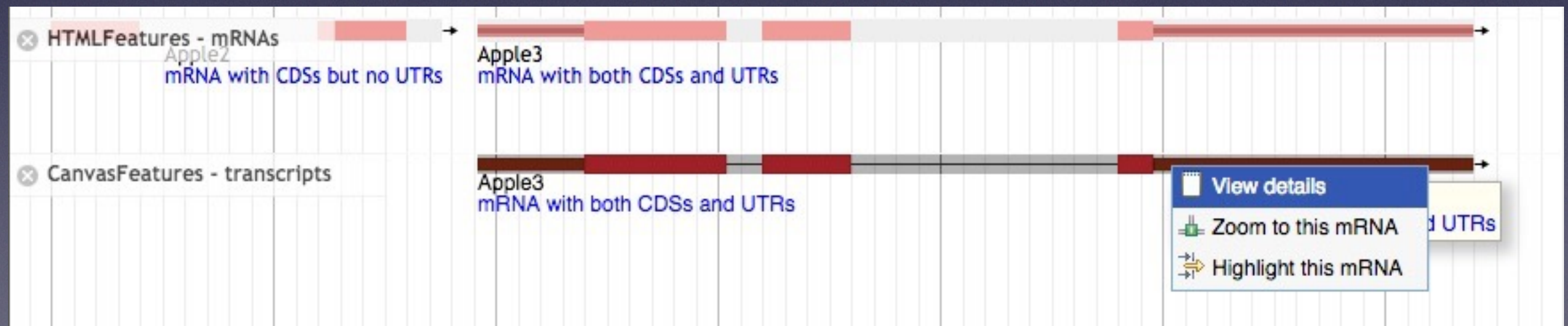
# The JBrowse Project

- free and open source (license: LGPL / Artistic)
- a GMOD project
  - <http://gmod.org>
- developed using git, hosted on GitHub
  - <http://github.com/GMOD/jbrowse>
- Pls most involved: Ian Holmes, Lincoln Stein, Suzi Lewis



# Feature Tracks

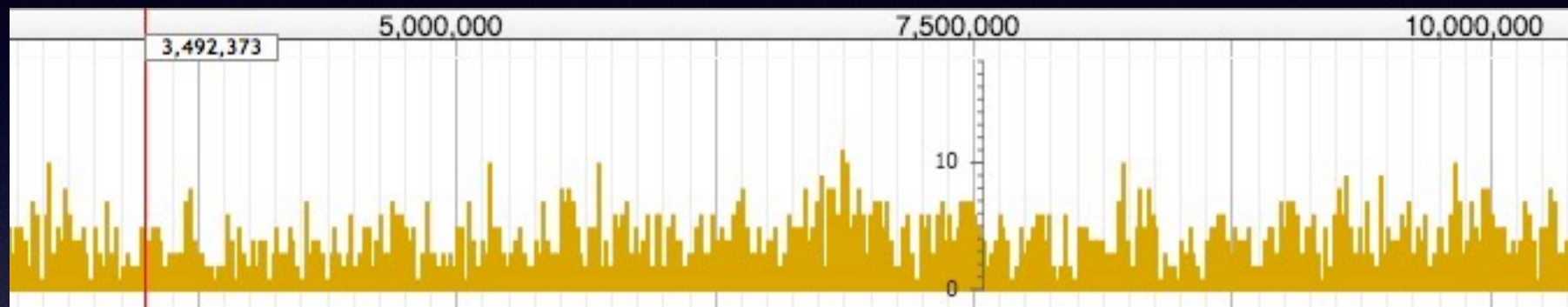
- HTMLFeatures: Rectangles (<div>s) with various fills and heights to represent the feature spans (more compatible)
- CanvasFeatures: Much prettier, more configurable glyphs
- Super-configurable left clicking and right-click



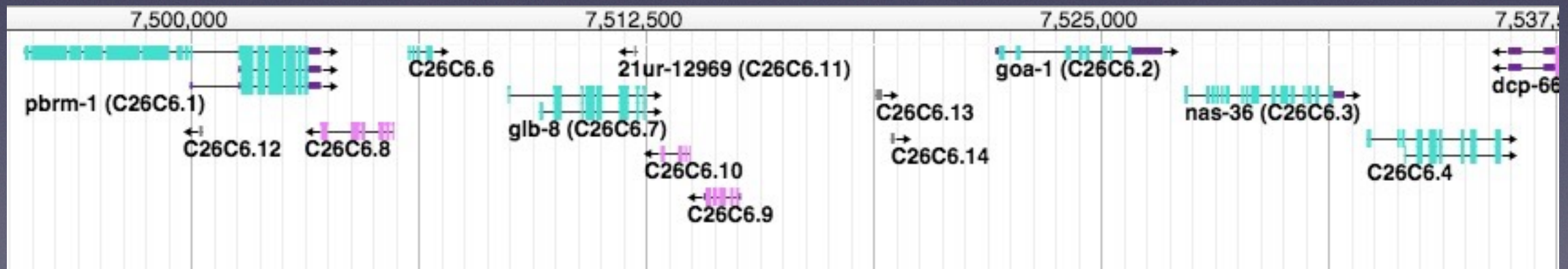


# Feature Density Plots

Zoomed out

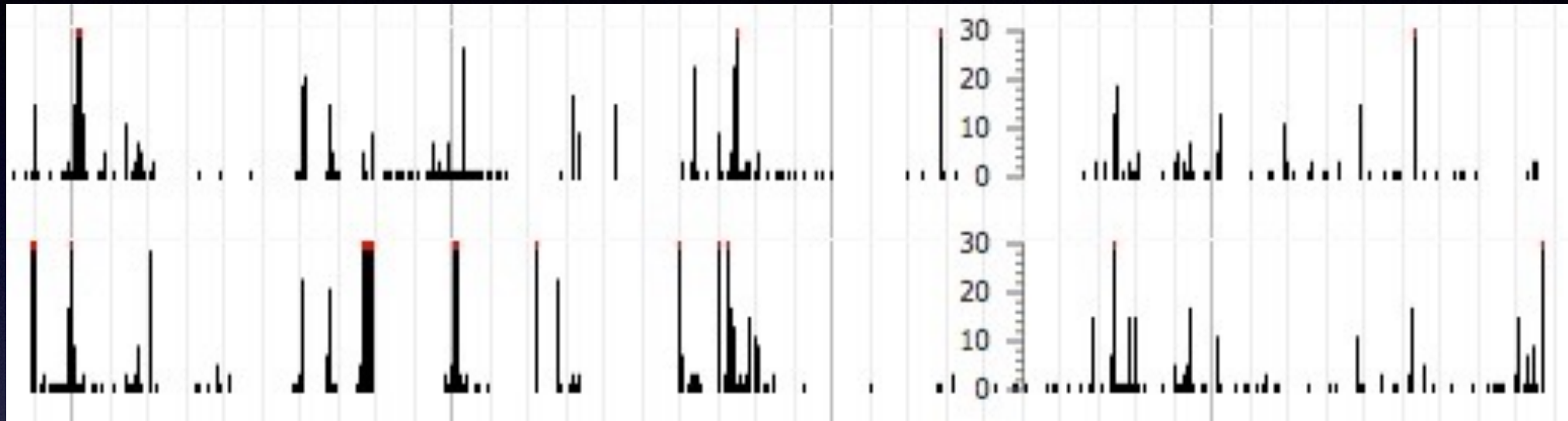


Zoomed in





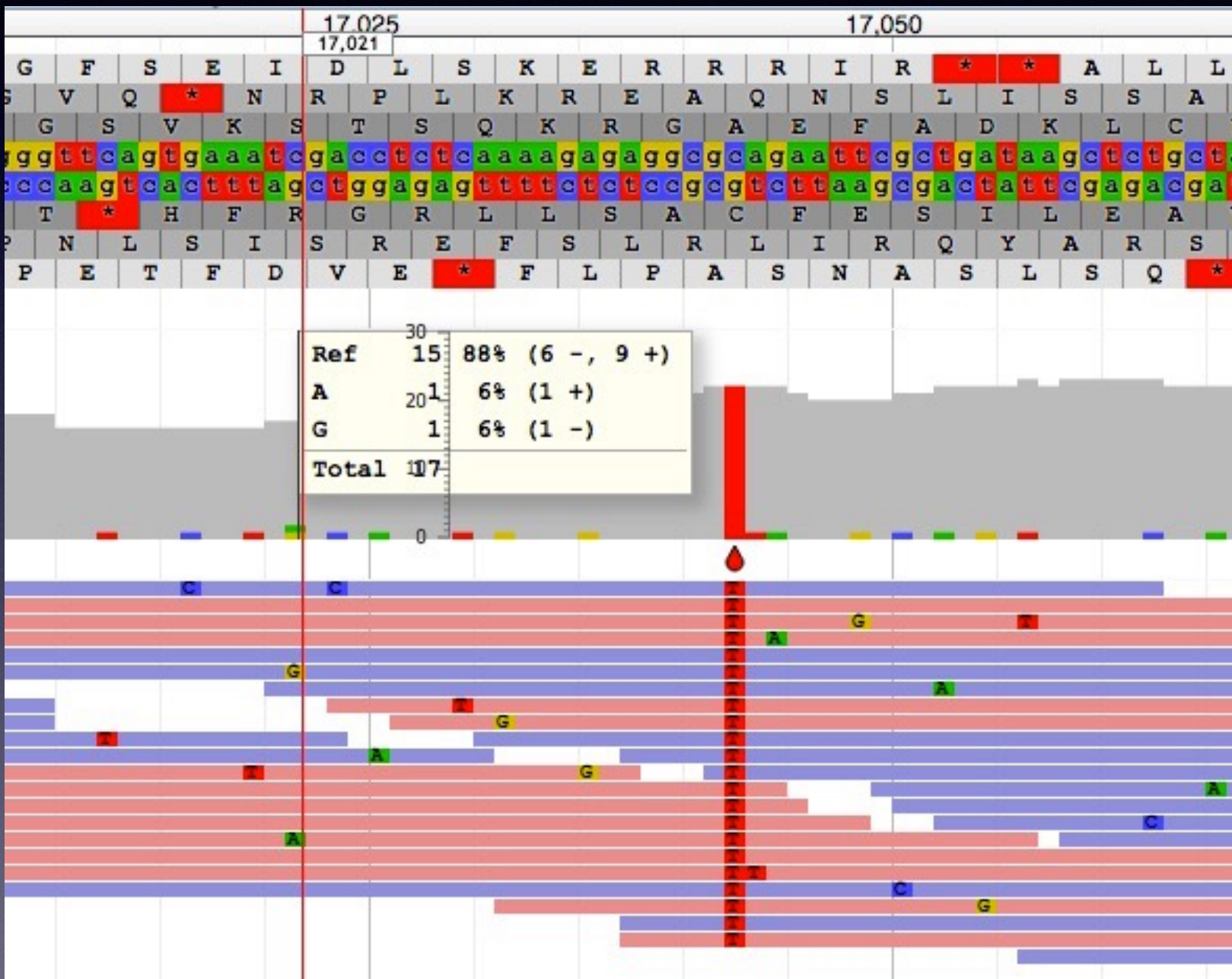
# Wiggle/BigWig Tracks



- Reads small chunks directly from BigWig file.
- Needs only a not-super-old ( $< 5$  yrs) browser, except for Internet Explorer.
- IE requires version 10.



# BAM Alignment Tracks

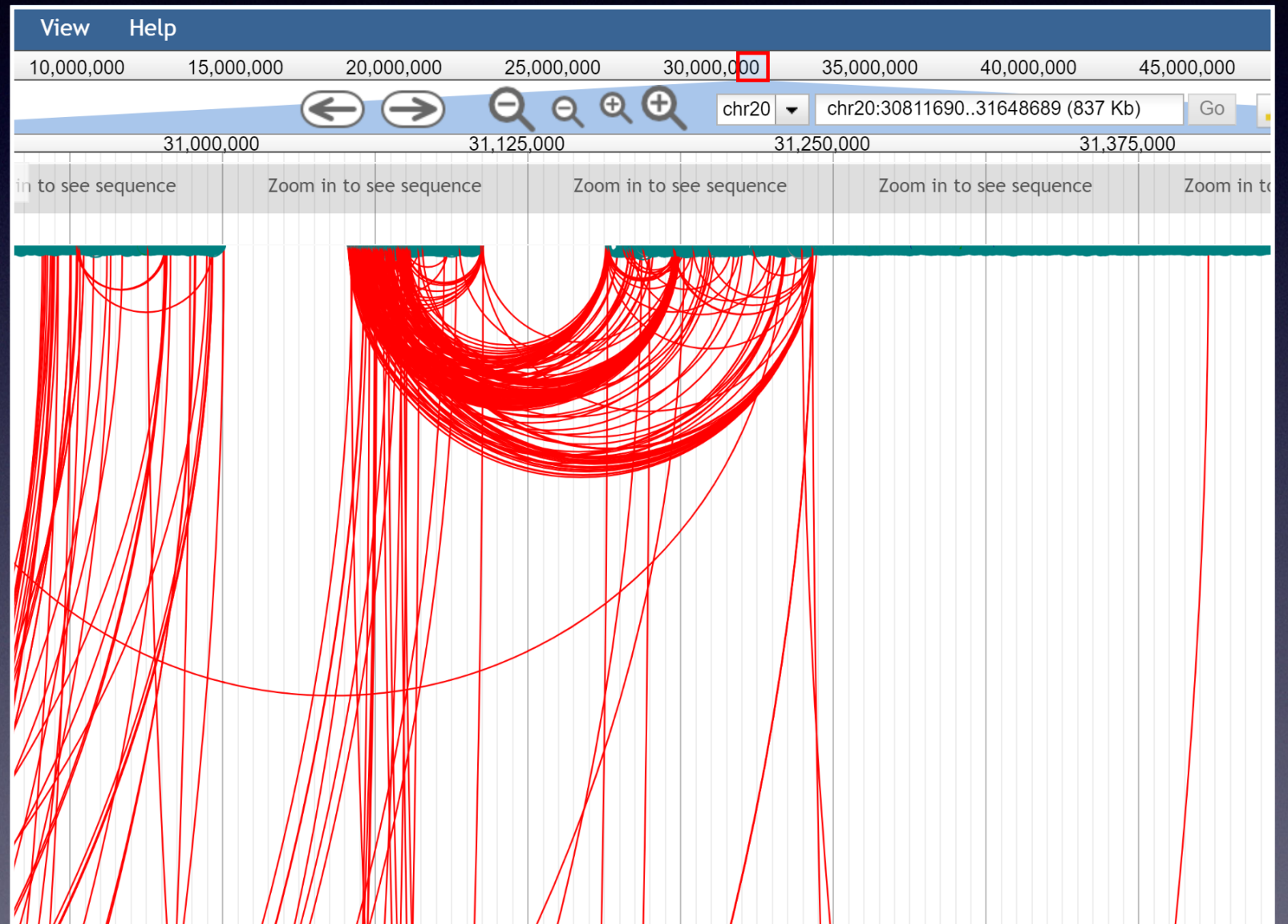


- Reads small chunks directly from BAM file.
- Coverage and mismatches.



# CRAM tracks too

- Better compression than BAM
- Paired reads (tho BAM has those now too)
- Lots of display options





# VCF Tracks

SNV

Primary Data

Type	SNV
Score	104
Description	SNV A -> T
Position	scf1117875582023:88017..88017
Length	1 bp

Attributes

AC1	2
AF1	1
DP	17
DP4	0 0 11 6
FQ	-78
MQ	57
VDB	0.0507
alternative_alleles	T
description	SNV A -> T
reference_allele	A
seq_id	scf1117875582023

Genotypes (1)

variant	1	100%
homozygous	1	100%
T variant	1	100%
Total	1	100%

Name	GT	PL	GQ
simulated-sorted.bam	T / T	137 51 0	99

- Reads directly from VCF files compressed and indexed with bgzip and tabix.
- Shows all VCF data: alleles, genotypes, quality, etc.



# Particular Strengths

- Web-based, but fast and smooth easy to set up
- Compressed NGS data: direct-to-browser BAM, BigWig, and VCF
- Optional faceted track selector – efficiently search thousands of tracks
- Open local files directly on client, no data transfer required
- Highly customizable, embeddable, integratable, programmable



# Apollo

<http://gmod.org/wiki/WebApollo>

- Based on JBrowse, using plugin system
- Next generation of the popular Apollo annotation editor
- Online annotation editing and curation!





# Apollo

- Clients receive updates in real time (like Google Docs)
- Saves edits to a central Chado database
- Client side is a JBrowse plugin
- Extensive server-side Java
- Maybe a live demo (later)
- <http://genomearchitect.org/WebApolloDemo/>



# JBrowse Plugins

- Extend JBrowse with your own JavaScript code
- Can do pretty much anything
  - Add your own track visualizations
  - Add your own data backends
  - Add menu items
  - Subscribe to event notifications (pub/sub system)
  - Reach deep into the guts of JBrowse and (carefully!)
  - change anything at all!
- Apollo client is a JBrowse plugin



# Coming future releases

- MORE: data types, sorting options, speed
- Graphical configuration
- Multiple views, linked or independent
- Logins, uploading, track sharing
- Server side component, BLAST integration



# Big Thanks

## **Ian Holms (UC Berkeley)**

Rob Buels

Mitch Skinner

Amelia Ireland

Eric Yao

Colin Diesh

## **Suzi Lewis (LBNL)**

Gregg Helt

Ed Lee

Justin Reese (UofMo)

## **Lincoln Stein (OICR)**

Julien Smith-Roberge

Erik Derohanian

Julie Moon

Natalie Fox

Adam Wright

## **NHGRI**

Cold, hard cash



# The End (on to the workshop)

<http://jbrowse.org/>

GMOD: <http://gmod.org/wiki/JBrowse>

Github: <http://github.com/GMOD/jbrowse>